
Using Puddle at H2O

Release Pre-release

H2O.ai

Mar 07, 2019

CONTENTS

1	Register for an Account	3
2	Starting Puddle at H2O	5
2.1	Puddle Account Options	5
3	Create New Driverless AI System	7
3.1	Viewing Driverless AI System Information	8
3.2	Starting Driverless AI	11
3.3	Launching Jupyter Notebook	12
3.4	SSH into the Driverless AI System	12
3.5	Stopping a System	12
3.6	Starting a System	13
3.7	Rebooting a System	13
3.8	Deleting a System	13
3.9	Editing an Annotation	13
3.10	Request Support	13
4	Create New H2O-3 System	15
4.1	Viewing H2O-3 System Information	16
4.2	Starting H2O-3	18
4.3	SSH into the H2O-3 System	20
4.4	Stopping a System	20
4.5	Starting a System	20
4.6	Rebooting a System	20
4.7	Deleting a System	21
4.8	Editing an Annotation	21
4.9	Request Support	21
5	FAQ	23

Welcome to Puddle at H2O, H2O.ai's software-as-a-service platform for running Driverless AI and H2O-3. Puddle is recommended for use as a POC, a demo, or for users who bring their own data.

Puddle is a single user platform. Accounts are tied to each user's IP. This means that multiple users cannot sign in to the same Puddle instance.

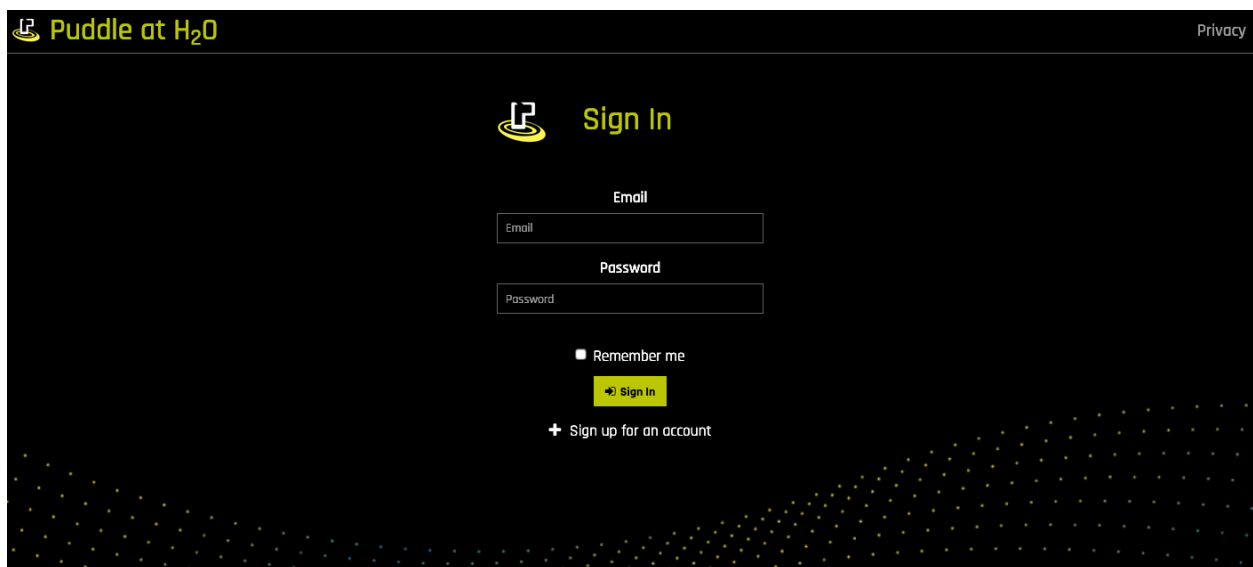
This is currently a Pre-release version of Puddle and, as such, is subject to change at frequent intervals.

Have Questions?

If you have questions about using Puddle, post them on the [H2O.ai Community Slack workspace](#) in the **#cloud** channel. If you have not signed up for the H2O.ai Community Slack workspace, you can do so here: <https://www.h2o.ai/community/>.

REGISTER FOR AN ACCOUNT

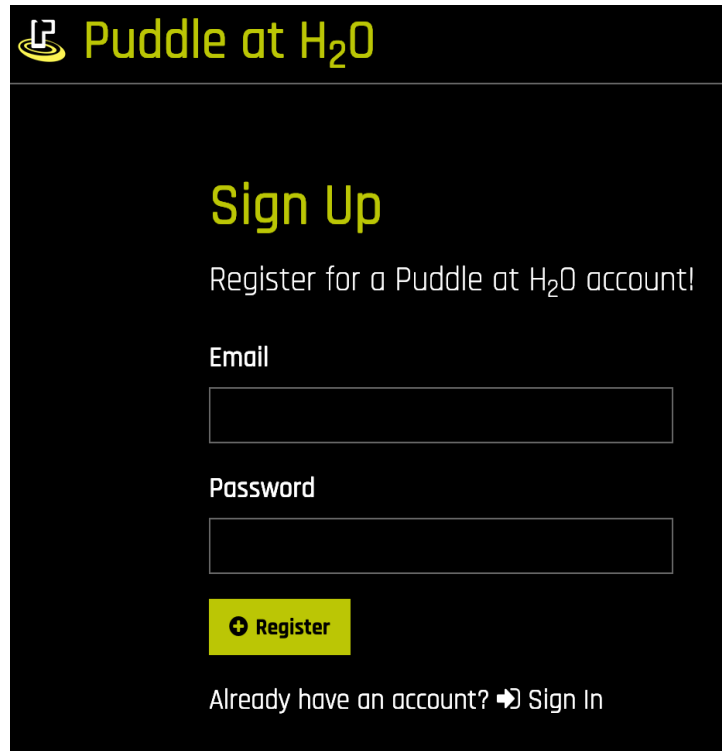
1. Go to puddle.h2o.ai and click **Sign Up for Account**.



The screenshot shows the 'Sign In' page for 'Puddle at H2O'. The page has a dark background with a pattern of small yellow and blue dots at the bottom. At the top left, there is a logo for 'Puddle at H2O' and a 'Privacy' link at the top right. The main heading is 'Sign In' with a logo icon. Below the heading are two input fields: 'Email' and 'Password'. Under the 'Password' field is a checkbox labeled 'Remember me'. Below the checkbox is a yellow 'Sign In' button. At the bottom, there is a link that says '+ Sign up for an account'.

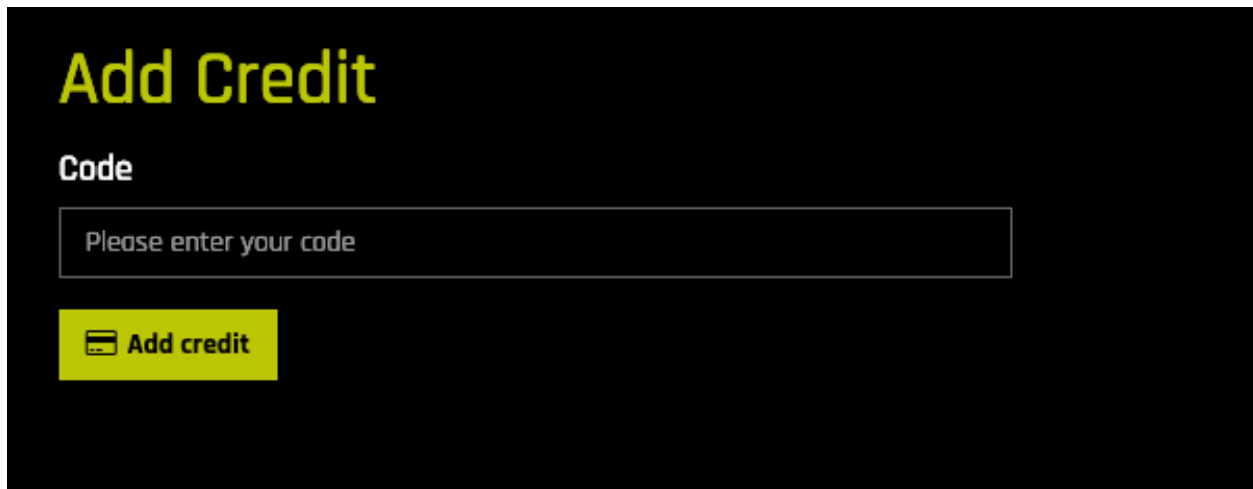
2. Enter your e-mail address and specify a password. Keep this information in a safe place. This will be your login information.

WARNING: You must contact your admin if you lose your password.

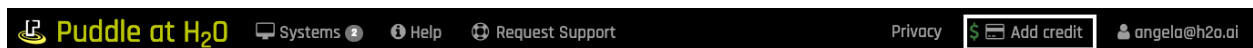


3. Click **Register**.

Upon completion, the Puddle at H2O site will launch. At this point, you will be prompted to provide a credit code. You will not be able to launch any systems until you apply credit. Enter your credit code, then click the **Add Credit** button to apply your credit.

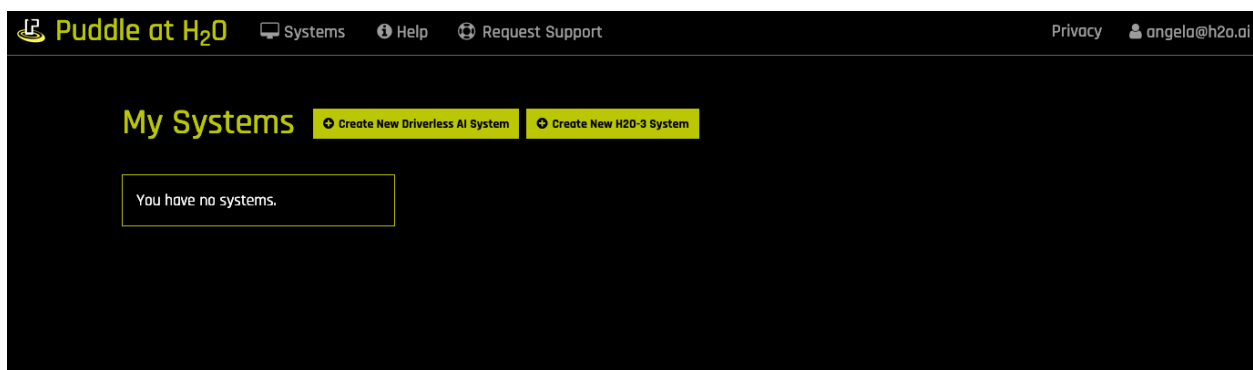


Note that you can also add credit at any time by clicking the **Add Credit** button in the top menu bar.



STARTING PUDDLE AT H2O

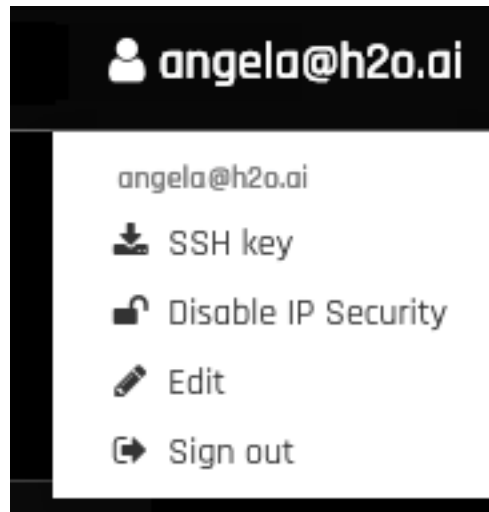
Sign in to your Puddle account at puddle.h2o.ai. The home page displays and shows any current systems. From this page, you can create a new Driverless AI and/or H2O-3 system.



2.1 Puddle Account Options

Clicking on your user name on the Puddle page provides several additional options:

- **SSH key:** Download the SSH key that can be used for accessing any running Driverless AI or H2O-3 systems.
- **Disable IP Security:** Puddle at H2O provides IP-based security. When you log in, Puddle retrieves your IP address and creates a personal security group that only allows your IP address to access it. Note that this means that if your IP address changes (for example, if you move your laptop to a different location resulting in an update to your IP address), then you must log in to Puddle again. This option allows you to disable IP security. Use caution when disabling IP security. Disabling security will allow connections from any IP address to all of your systems. If you select this option, you will be prompted to confirm the disable.
- **Edit:** This option allows you to change your password.
- **Sign out:** Signs you out of Puddle. **CAUTION:** Signing out does not stop running systems.



CREATE NEW DRIVERLESS AI SYSTEM

This section describes how to create a new Driverless AI system.

Note: You must have a Driverless AI license in order to run Driverless AI in Puddle. You can request a free 21-day trial license at <https://www.h2o.ai/try-driverless-ai/>.

1. Click **Create New Driverless AI System** on the Puddle Systems page.
2. Specify the following options to create the system:
 - Driverless AI Version: This defaults to the most current version of Driverless AI.
 - System Name: This is blank by default.
 - System Type: This can be CPU Small (default), CPU Large, GPU Small, GPU Memory, or GPU Compute. Please note the cost for each system type. **Note:** Your account settings may include a limit as to the number of systems of a certain type that you can run. If you exceed that limit (for example, if you exceed the number of CPU Small systems that may configure for Driverless AI), then that option will not be available, and the least expensive option will then become the default.
 - Volume Size: This can be 256GiB (default), 512GiB, or 1TiB.
 - Stopped If Idle For: This can be 30 min, 1 hour (default), 2 hours, 3 hours, or 4 hours.
 - Annotation: Specify a description for the system. This is blank by default. It can also be edited directly on the system page.

Create New System

Driverless AI Version
1.5.3

System Name
Name

System Type
CPU Small - \$0.38 per Hour

Volume Size
256GiB

Stopped If Idle For
1h

Annotation
Annotation

+ Create System

3. Click **Create System** when you are done. The system will begin initializing. Note that this can take several minutes.

After the system has completed initializing, you will be ready to use Driverless AI.

3.1 Viewing Driverless AI System Information

Click on the Driverless AI system to view the configuration information, model information for available experiments, and system metrics.

The screenshot shows the 'DAI Test' interface for Driverless AI. It features a top navigation bar with icons for refresh, delete, edit, and a globe. Below this is a list of system details:

- URL:** 35.173.245.47:12345 or 35.173.245.47
- Jupyter:** Open Jupyter Notebook
- Status:** Started
- Username:** h2oai
- Password:** i-01ceee67a5a42a684
- Configuration:** Edit config.toml
- SSH Key:** Get SSH Key
- SSH Command:** ssh -i puddle-key-angela@h2o.ai-production.pem ubuntu@35.173.245.47
- Stopped if Idle Till:** Wed Feb 20 2019 14:38:52 GMT-0800 (Pacific Standard Time) with a Reset timer button.
- Price for This Session:** \$0.38
- Price for Previous Sessions:** \$0.00
- Product:** DAI
- Version:** 1.5.4
- System Type:** CPU Small, \$0.38 per Hour
- Created:** Wed Feb 20 2019 13:33:33 GMT-0800 (Pacific Standard Time)

At the bottom, there is a 'DriverlessAI Models' section with a table of metrics:

Key	Dataset	Target	Validation Score	Test Scorer	Scorers	Progress	Status	Accuracy	Time	Interpretability	Training Duration
-----	---------	--------	------------------	-------------	---------	----------	--------	----------	------	------------------	-------------------

This page provides general system information, Driverless AI model information (if any models exist), and System Metrics graphs.

3.1.1 System Information

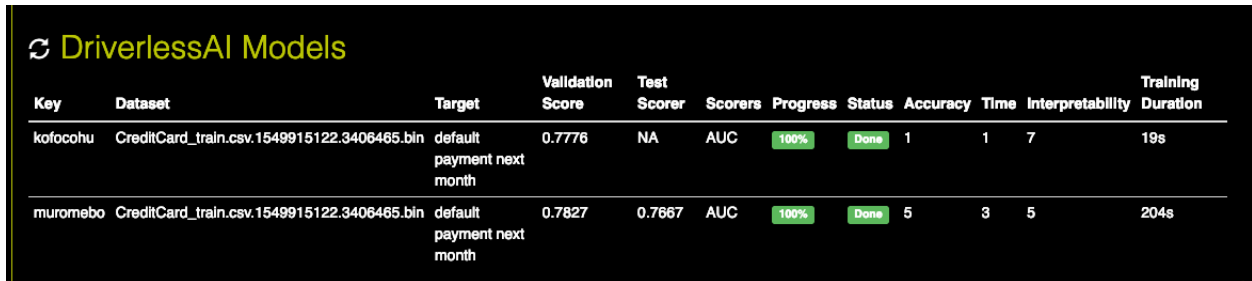
- The URL for starting Driverless AI (with and without port value). Clicking this takes you to the DNS of the URL.
- A link to open a Jupyter Notebook. This can then be used for running a Driverless AI experiment. (See *Launching Jupyter Notebook* below for more info.)
- The system status.
- The Username and Password for starting Driverless AI.
- A link for editing the config.toml file. The config.toml file includes all possible configuration options that can be specified when starting Driverless AI. A list of available configuration options is available in the *Driverless AI User Guide*.
- An SSH key for accessing the Driverless AI system. Download this key to enable secure access to this Driverless AI system. If necessary, run `chmod 400 puddle-key-<puddle_email>-production.pem` to ensure that your private key file is not accessible by others.
- The SSH command to run in order to securely access the system that is running Driverless AI. Be sure to update this path to point to where you stored the downloaded SSH key. (See *SSH into the Driverless AI System* below for more information.)
- The time when the system will stop if remaining idle. You can also reset this timer.
- The price for the current session and any previous sessions.
- The product name and version currently running on the system.

- The system type along with the cost per hour.
- The creation date.

3.1.2 Driverless AI Model Information

For each experiment run on Driverless AI through Puddle, the following information displays:

- The experiment key.
- The training dataset used in the experiment.
- The target column.
- The validation score.
- The test scorer (if a validation set is not provided).
- The scorer used for the experiment.
- The experiment progress and status.
- The Accuracy, Time, and Interpretability options used for the experiment.
- The time it took to complete the experiment (in seconds).



Key	Dataset	Target	Validation Score	Test Scorer	Scorers	Progress	Status	Accuracy	Time	Interpretability	Training Duration
kofocohu	CreditCard_train.csv.1549915122.3406465.bin	default payment next month	0.7776	NA	AUC	100%	Done	1	1	7	19s
muromebo	CreditCard_train.csv.1549915122.3406465.bin	default payment next month	0.7827	0.7867	AUC	100%	Done	5	3	5	204s

3.1.3 System Metrics

For each system, the following graphs are provided:

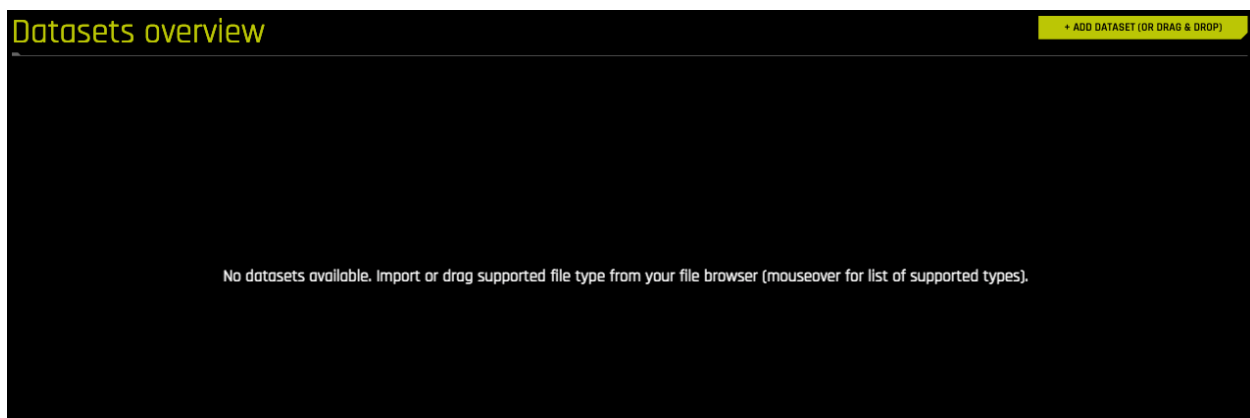
- CPU/GPU utilization graph.
- CPU/GPU credit usage (if credit was available).
- Disk Read.
- Disk Write.
- Network In.
- Network Out.



3.2 Starting Driverless AI

1. Click on the URL provided in the Driverless AI system page. This takes you to the DNS of the URL.
2. If this is your first time starting Driverless AI on this system, or if you have restarted the system, accept the license agreement.
3. Enter the Username and Password that are provided on the Driverless AI system page.
4. If this is your first time starting Driverless AI on this system, enter your license key. Note that if you do not have a license key, you can request a free 21-day trial license at <https://www.h2o.ai/try-driverless-ai/>.

Upon completion, Driverless AI will open on the Datasets Overview page. At this point, you can add or upload datasets and begin running experiments.



In Driverless AI, click on **Resources > Help** to view the Driverless AI User Guide. Additional documentation for Driverless AI is available at docs.h2o.ai.

3.3 Launching Jupyter Notebook

Puddle includes a link for quickly opening Jupyter Notebook. This can then be used for running a Driverless AI experiment using the Python client.

When you launch Jupyter Notebook from Puddle, you will see the following folders:

- **data:** This includes sample credit card data for running an experiment through Jupyter Notebook.
- **h2oai_client_demo:** This includes an example Notebook showing how to use the Python API in Jupyter Notebook to add datasets, run and examine an experiment, and download predictions.
- **notebooks:** A folder for storing additional Jupyter Notebooks.

3.4 SSH into the Driverless AI System

Puddle provides the ability to SSH into a system that is running Driverless AI.

1. Select the system that you want to SSH into.
2. Download the SSH key by clicking the **Get SSH Key** link. Save this on your local machine.
3. On your local machine, run `chmod 400 puddle-key-<puddle_email>-production.pem` to ensure that your private key file is not accessible by others. (Replace `<puddle_email>` in the command with your Puddle at H2O email address.)
4. On your local machine, run the provided SSH command. If necessary, update this path to point to where you stored the downloaded SSH key.


After you are SSH'd into the Driverless AI system, you will be able to access logs, datasets, and all data stored in the tmp folder, or you can copy data into that system using scp. For example:

```
# SSH into the DAI machine
ssh -i puddle-key-angela@h2o.ai-production.pem ubuntu@35.173.245.47


# Review files and folders on the DAI machine
ubuntu@ip-10-10-1-53:~$ ls
config                data                  license  tmp
dai-docker-centos7-x86_64-1.5.3-9.0.tar.gz  demo                 log
DAI_RELEASE           jupyter              scripts

# Open another terminal window and copy data from your
# local machine onto the DAI machine's /data folder.
# Replace the path for your pem key and dataset folder.
$ scp -i "puddle-key-angela@h2o.ai-production.pem" ~/datasets/CreditCard_test.csv_
↪ubuntu@35.173.245.47:/data
```


3.5 Stopping a System

Click the Stop button () to halt the system that is in a “Started” state. No information will be lost when a system is stopped.


3.6 Starting a System

Click on the Start button () to start a system that is in a “Stopped” state. This will launch a new system with a new URL. All prior data will still be available from this URL.


3.7 Rebooting a System

Click the Reboot button () to reboot a system that is in a “Started” state. This will stop the system and launch a new system with a new URL. All prior data will still be available from the new URL. The entire process can take several minutes.


3.8 Deleting a System

Click the Delete button () to completely remove a system. A confirmation page will display asking if you are certain about deleting the system. Click **Yes** complete the delete. This request deletes the system and destroys all data that is on the system.

3.9 Editing an Annotation


Click the Edit button () to change the annotation of a system, then click **Update System** to complete the change.

3.10 Request Support

Click the Request Support button () or click the **Request Support** link on the top menu to send any issues to support. This opens the help form. Enter your issue in the form, and then click **Send** when you are done or click **Cancel** to return to the My Systems page. Responses to support requests will be sent to the e-mail address associated with your Puddle account. You can also follow up with your Sales Engagement contact.

Help

Please describe your problem



? Send

⊘ Cancel

CREATE NEW H2O-3 SYSTEM

This section describes how to create a new H2O-3 system. H2O-3 is open source and does not require a license.

1. Click **Create New H2O-3 System** on the Puddle Systems page.
2. Specify the following options to create the system:
 - H2O-3 Version: This defaults to the most current version of H2O-3.
 - System Name: This is blank by default.
 - System Type: This can be CPU Small (default), CPU Large, GPU Small, GPU Memory, or GPU Compute. Please note the cost for each system type. **Note:** Your account settings may include a limit as to the number of systems of a certain type that you can run. If you exceed that limit (for example, if you exceed the number of CPU Small systems that may configure for H2O-3), then that option will not be available, and the least expensive option will then become the default.
 - Volume Size: This can be 256GiB (default), 512GiB, or 1TiB.
 - Stopped If Idle For: This can be 30 min, 1 hour (default), 2 hours, 3 hours, or 4 hours.
 - Annotation: This is blank by default.

Create New System

H2O-3 Version

3.22.1.3

System Name

Name

System Type

CPU Small - \$0.80 per Hour

Volume Size

256GiB

Stopped If Idle For

1h

Annotation

Annotation

Create System

3. Click **Create System** when you are done. The system will begin initializing. Note that this can take several minutes.

4.1 Viewing H2O-3 System Information

Click on the Driverless AI system to view its configuration information and system metrics.

H2O-3 Test Test H2O-3 in Puddle

URL: 54.209.253.109:54321

Status: **Started**

SSH Key: [Get SSH Key](#)

SSH Command: `ssh -i puddle-key-angela@h2o.ai-production.pem ubuntu@54.209.253.109`

Stopped If Idle Till: Wed Feb 20 2019 14:40:42 GMT-0800 (Pacific Standard Time) [Reset timer](#)

Price for This Session: \$0.80

Price for Previous Sessions: \$0.00

Product: **H2O-3**

Version: **3.22.1.3**

System Type: **CPU Small** \$0.80 per Hour

Created: Wed Feb 20 2019 13:38:19 GMT-0800 (Pacific Standard Time)

System Metrics

CPU Utilization

Time	CPU Utilization (%)
11:00	0.637
13:30	1.27

CPU Credit Usage

Time	CPU Credit Usage
11:00	0.0
13:30	0.0

Disk Read

Time	Disk Read (Bytes)
11:00	0.0
13:30	0.0

This page provides general system information as well as System Metrics graphs.

4.1.1 System Information

- The URL for launching H2O-3. Click on the URL to open Flow. This takes you to the DNS of the URL. **Note:** You can also download the H2O-3 Python or R clients from <https://www.h2o.ai/download/>, and use this URL to launch H2O-3 in either Python or R. But be sure that the H2O-3 version in Puddle and the version downloaded locally match, or you will receive a version mismatch error.
- The system status.
- An SSH key for accessing the H2O-3 system. Download this key to enable secure access to this H2O-3 system. If necessary, run `chmod 400 puddle-key-<puddle_email>-production.pem` to ensure that your private key file is not accessible by others.
- The SSH command to run in order to securely access the system that is running Driverless AI. Be sure to update this path to point to where you stored the downloaded SSH key. (See *SSH into the H2O-3 System* below for more information.)
- The time when the system will stop if remaining idle. You can also reset this timer.
- The price for the current session and any previous sessions.
- The product name and version currently running on the system.
- The system type along with the cost per hour.
- The creation date.

4.1.2 System Metrics

For each system, the following graphs are provided:

- CPU/GPU utilization graph.
- CPU/GPU credit usage (if credit was available).
- Disk Read.
- Disk Write.
- Network In.
- Network Out.



4.2 Starting H2O-3

After the system has completed initializing, you will be ready to use H2O-3. As indicated previously, H2O-3 can be started in Flow, Python, or R.

4.2.1 Starting H2O-3 in Flow

Click on the URL to launch H2O Flow. This also takes you to the DNS of the URL. Refer to the Flow section in the [H2O-3 User Guide](#) for information on how to use H2O-3.

4.2.2 Starting the H2O-3 Python Client

Perform the following steps to start the H2O-3 Python client. Run these commands in a Terminal window.

Note: These steps assume that the H2O-3 system created in Puddle is using the latest H2O-3 release. If your H2O-3 version in Puddle is not the latest, then refer to the [H2O-3 Prior Releases](#) page for information on downloading the H2O-3 Python client for your specific version.

1. Install dependencies (prepending with `sudo` if needed):

```
pip install requests
pip install tabulate
pip install "colorama>=0.3.8"
pip install future
```

Note: These are the dependencies required to run H2O. A complete list of dependencies is maintained in the following file: <https://github.com/h2oai/h2o-3/blob/master/h2o-py/conda/h2o/meta.yaml>.

2. Run the following command to remove any existing H2O module for Python.

```
pip uninstall h2o
```

3. Use `pip` to install this version of the H2O Python module.

```
pip install -f http://h2o-release.s3.amazonaws.com/h2o/latest_stable_Py.html_
↪h2o
```

Note: When installing H2O from `pip` in OS X El Capitan, users must include the `--user` flag. For example:

```
pip install -f http://h2o-release.s3.amazonaws.com/h2o/latest_stable_Py.html_
↪h2o --user
```

4. Initialize H2O in Python and optionally run a demo to see H2O at work. In the `h2o.init()` command, be sure to change `<h2o_3_system_address>` to the URL provided in Puddle for this H2O-3 system.

```
python
import h2o
h2o.init("ip=<h2o_3_system_address>", port=54321)
h2o.demo("glm")
```

4.2.3 Starting the H2O-3 R Client

Perform the following steps to start the H2O-3 Python client. Run the commands in R, copying and pasting these commands one line at a time.

Note: These steps assume that the H2O-3 system created in Puddle is using the latest H2O-3 release. If your H2O-3 version in Puddle is not the latest, then refer to the [H2O-3 Prior Releases](#) page for information on downloading the H2O-3 R client for your specific version.

1. The following two commands remove any previously installed H2O packages for R.

```
if ("package:h2o" %in% search()) { detach("package:h2o", unload=TRUE) }
if ("h2o" %in% rownames(installed.packages())) { remove.packages("h2o") }
```

2. Next, download packages that H2O depends on.

```
pkgs <- c("RCurl", "jsonlite")
for (pkg in pkgs) {
  if (!(pkg %in% rownames(installed.packages()))) { install.packages(pkg) }
}
```

3. Download and install the H2O package for R.

```
install.packages("h2o", type="source", repos=(c("http://h2o-release.s3.  
→amazonaws.com/h2o/latest_stable_R")))
```

4. Initialize H2O and optionally run a demo to see H2O at work. In the `h2o.init()` command, be sure to change `<h2o_3_system_address>` to the URL provided in Puddle for this H2O-3 system.


```
library(h2o)  
h2o.init("ip=<h2o_3_system_address>", port=54321)  
demo(h2o.kmeans)
```

4.3 SSH into the H2O-3 System


Puddle provides the ability to SSH into a system that is running H2O-3.

1. Select the system that you want to SSH into.
2. Download the SSH key by clicking the **Get SSH Key** link. Save this on your local machine.
3. On your local machine, run `chmod 400 puddle-key-<puddle_email>-production.pem` to ensure that your private key file is not accessible by others. (Replace `<puddle_email>` in the command with your Puddle at H2O email address.)
4. On your local machine, run the provided SSH command. If necessary, update this path to point to where you stored the downloaded SSH key.


4.4 Stopping a System

Click the Stop button () to halt the system that is in a “Started” state. No information will be lost when a system is stopped.


4.5 Starting a System

Click on the Start button () to start a system that is in a “Stopped” state. This will launch a new system with a new URL. All prior data will still be available from this URL.


4.6 Rebooting a System

Click the Reboot button () to reboot a system that is in a “Started” state. This will stop the system and launch a new system with a new URL. All prior data will still be available from the new URL. The entire process can take several minutes.


4.7 Deleting a System

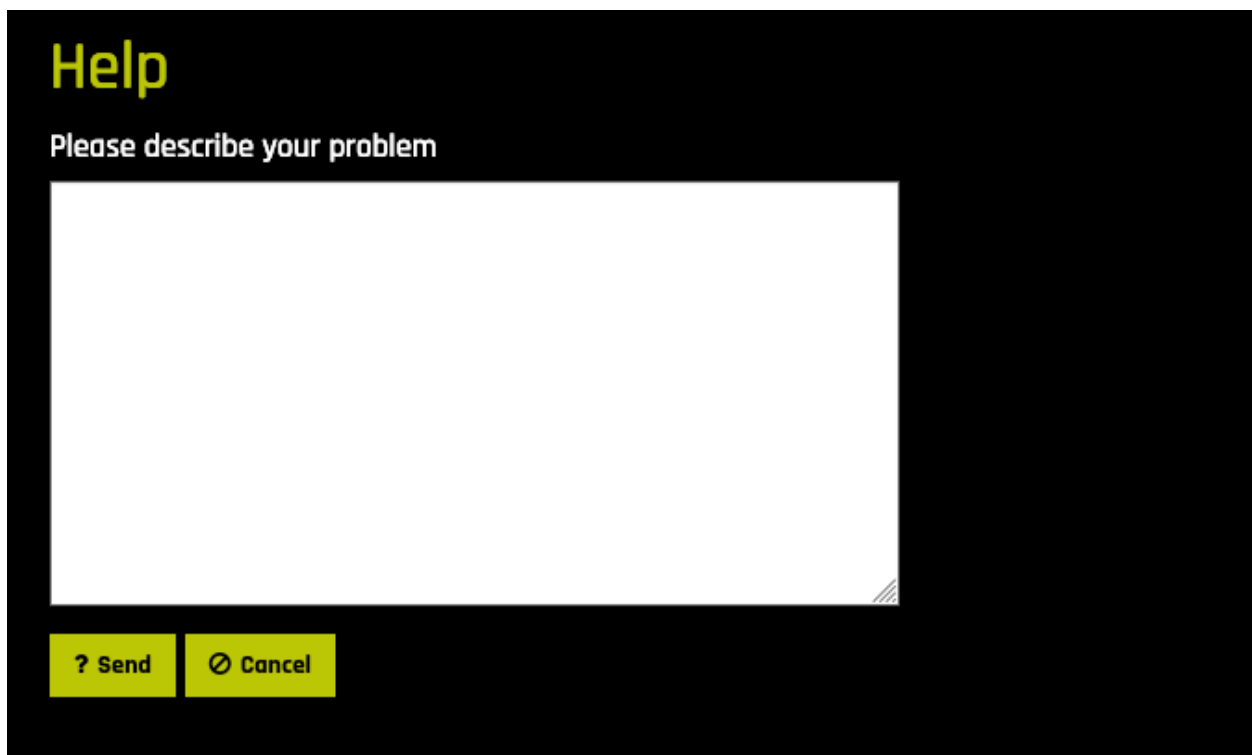
Click the Delete button () to completely remove a system. A confirmation page will display asking if you are certain about deleting the system. Click **Yes** complete the delete. This request deletes the system and destroys all data that is on the system.

4.8 Editing an Annotation

Click the Edit button () to change the annotation of a system, then click **Update System** to complete the change.

4.9 Request Support

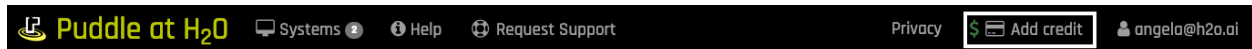
Click the Request Support button () or click the **Request Support** link on the top menu to send any issues to support. This opens the help form. Enter your issue in the form, and then click **Send** when you are done or click **Cancel** to return to the My Systems page. Responses to support requests will be sent to the e-mail address associated with your Puddle account. You can also follow up with your Sales Engagement contact.



The screenshot shows a dark-themed help form. At the top left, the word "Help" is written in a large, yellow font. Below it, the text "Please describe your problem" is displayed in a smaller, white font. A large, empty white rectangular area occupies the center of the form, intended for the user to type their issue. At the bottom of the form, there are two yellow buttons: one with a question mark icon and the text "? Send", and another with a circular arrow icon and the text "Cancel".

I have a credit code available for my Puddle account. How can I apply this credit?

If you have credit available for your Puddle account, then an **Add Credit** button will be available on the top menu bar.



Click that button to open the Add Credit form. Enter your credit code, then click the **Add Credit** button to apply your credit.



How can I recover my password?

Currently, there is no way to recover a password. You must request a new one from the admin.

What type of security is enabled in Puddle?

Puddle at H2O provides IP-based security. When you log in, Puddle retrieves your IP address and creates a personal security group that only allows your IP address to access it. Note that this means that if your IP changes (for example, if you move your laptop to a different location resulting in an update to your IP), then you must log in to Puddle again. Though not recommended, you can turn off IP security by selecting **Disable IP Security** in your account options.

Is data encrypted in cloud storage?

No, data is not encrypted.

Are data transfers encrypted and secured?

No, data transfers are not encrypted/secured.

Where is my data stored?

All data is stored on your AWS instance. Keep in mind that if you terminate your instance, all data will be lost and cannot be retrieved. Note that though H2O owns the instance, Puddle's IP-based security prevents us from accessing any instance or retrieving/accessing any data on an instance.

Can I delete data and models from the Driverless AI or H2O-3 systems?

Models can be deleted directly from within Driverless AI. Also, terminating/deleting a system destroys all data in a Driverless AI or H2O-3 system, including datasets and models/experiments.

Why am I seeing a "Puddle is currently experiencing high demand" message?

There may be times when you receive the following message: "Failed - Puddle is currently experiencing high demand and there is no machine capacity of this type available, please try again later, or attempt to use a different machine type." This message is due to AWS running out of resources.